

# Transposition and targeting of the prokaryotic mobile element IS30 in zebrafish

Mónika Szabó<sup>a,1</sup>, Ferenc Müller<sup>b,c,1</sup>, János Kiss<sup>a</sup>, Carolin Balduf<sup>b</sup>, Uwe Strähle<sup>c,\*</sup>,  
Ferenc Olsz<sup>a,\*\*</sup>

<sup>a</sup>Environmental Biosafety Research Institute, Agricultural Biotechnology Center, Szent-Györgyi Albert St. 4, H-2101 Gödöllő, Hungary

<sup>b</sup>Institute of Toxicology and Genetics, Forschungszentrum, Karlsruhe, Postfach 3640, D-76021 Karlsruhe, Germany

<sup>c</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, BP 10142, 67404 Illkirch Cedex, C.U. de Strasbourg, France

Received 2 June 2003; revised 7 July 2003; accepted 10 July 2003

First published online 30 July 2003

Edited by Giovanni Cesareni

**Abstract** We provide evidence that a prokaryotic insertion sequence (IS) element is active in a vertebrate system. The transposase of *Escherichia coli* element IS30 catalyzes both excision and integration in extrachromosomal DNA in zebrafish embryos. The transposase has a pronounced target preference, which is shown to be modified by fusing the enzyme to unrelated DNA binding proteins. Joining the transposase to the cI repressor of phage  $\lambda$  causes transposition primarily into the vicinity of the  $\lambda$  operator in *E. coli*, and linking to the DNA binding domain of Gli1 also directs the recombination activity of transposase near to the Gli1 binding site in zebrafish. Our results demonstrate the possibility of fusion transposases to acquire novel target specificity in both prokaryotes and eukaryotes.

© 2003 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.

**Key words:** IS30 transposase; Gli1;  $\lambda$  cI repressor; Protein fusion; Zebrafish; Site-directed integration

## 1. Introduction

Insertion sequences (IS) are mobile DNA segments possessing the ability to insert into the genome of host organisms. The well-characterized IS30 element of *Escherichia coli* flanked by 26 bp terminal inverted repeats (IRs) encodes the 44.3 kDa Tase. The Tase interacts with the IRs and does not appear to require *E. coli*-specific host factors. During transposition, an intermediate, (IS30)<sub>2</sub>, composed of two abutted IS copies that are separated by 2 bp spacer is formed. The joined IRs act as a recombinogenic site and are ca. 1000 times more efficient in transposition than the single element [1,2]. The resolution of (IS30)<sub>2</sub> during transposition results in a stable integration product. Several transposons were previously identified and isolated from fish such as *sleeping beauty* from salmonids [3] or the Tol2 element from medaka [4];

however, these systems were burdened by the restricted size of donor DNA that can be inserted by these ‘cut and paste’ transposases. In contrast, donor DNA of any theoretical size that harbors (IS30)<sub>2</sub> can be integrated by the IS30 Tase. Transposition of the *Drosophila* transposon *mariner* [5] in the protozoan *Leishmania* shows that transposons may retain activity in phylogenetically very distant species. Here, we investigated whether mobile elements can cross the prokaryotic/eukaryotic barrier by analyzing IS30 transpositional activity in vertebrate cells in plasmid-based systems.

To modify the pronounced site preference of the IS30 Tase [6], it was joined to heterologous DNA binding domains (DBD). The ability of chimeric proteins was assessed to direct integration into the proximity of target sites of the linked DBDs in *E. coli* and zebrafish. Our results provide the proof-of-principle for targeting recombinations using fusion-Tase proteins in living cells.

## 2. Materials and methods

### 2.1. Plasmids

The tester plasmid pJKI216 is described in [7]. The pCS2+-based [8] transposase producer constructs express IS30 Tase or its derivatives fused to the 6-mer of Myc-epitope tag (MT) or a nuclear localization signal (NLS) peptide or both. The circular *gfp-fg* donor (Fig. 1) was generated by ligation of the purified 2.6 kb *Bam*HI fragment of *gfp-p* (pMSZ198), a donor plasmid, where the *gfp* reporter gene was preceded by the 111 bp splice-acceptor sequence *intA* [9], an (IS30)<sub>2</sub> site and the Cm<sup>R</sup> bacterial marker. The *shh* target plasmids carry the 11.5 kb fragment of the zebrafish *sonic hedgehog* (*shh*) locus [10]. Target plasmids *shh-GOHS* and *shh-gli* were generated by inserting the 24 bp hot spot target sequence of IS30 (GOHS) [11] or the consensus Gli binding site (GACCACCCA, [12] into the *Stu*I site of *shh*, respectively. The Tase-cI chimeric protein was expressed from a p15A-based Km<sup>R</sup> vector containing the ORF-A of IS30 linked by the amino acids Leu and Gln C-terminally to the cI repressor gene of  $\lambda$  ts587 under the control of the *tac* promoter. The expression plasmid also harbors the (IS30)<sub>2</sub> structure and the Cm<sup>R</sup> gene. In control experiments, wild-type (wt) Tase was expressed from a similar construct without the cI repressor gene. Target plasmids were pEMBL19 and its derivatives containing GOHS or the 200 bp  $\lambda$  operator O<sub>R</sub> ( $\lambda$  coordinates: 37946–38107 bp, EM\_PH:M17233). The Tase-GliDBD fusion protein (expressed from a pCS2+-derived plasmid) contained the IS30 Tase linked to the N-terminus of the 256–417 aa region of the human Gli protein [12] via Glu and Phe.

### 2.2. DNA/RNA procedures and microinjection

DNA techniques were performed according to [13]. Sequences of polymerase chain reaction (PCR) primers are available upon request. *E. coli* strains TG2 and TOP10 Electrocomp<sup>®</sup> (Invitrogen) were used for cloning and electroporation of DNA isolated from fish embryos. Tase mRNA variants were synthesized in vitro by an Ambion

\*Corresponding author. Fax: (33)-3-88 65 32 01.

\*\*Also corresponding author. Fax: (36)-28-526 198.

E-mail addresses: [uwe@igbmc.u-strasbg.fr](mailto:uwe@igbmc.u-strasbg.fr) (U. Strähle), [olasz@abc.hu](mailto:olasz@abc.hu) (F. Olsz).

<sup>1</sup> These authors contributed equally to this work.

**Abbreviations:** DBD, DNA binding domain; GOHS, oligonucleotide hot spot based on *E. coli* genomic IS30 insertion sites; IS, insertion sequence; IR, inverted repeat; MT, Myc-epitope tag; NLS, nuclear localization signal

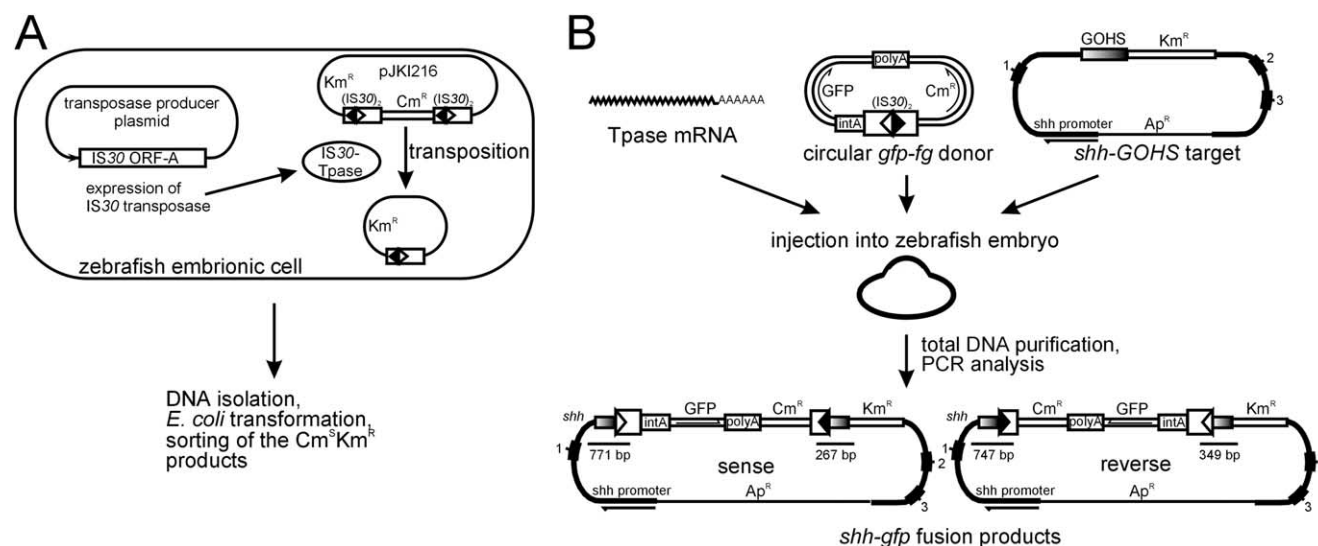


Fig. 1. Schematic representation of the experimental design used to detect transposition. A: The excision reaction generated by IS30 T<sub>pase</sub> in the tester plasmid. B: The experimental design used to detect insertions in *shh* target plasmids. In both experiments DNA or in vitro transcribed mRNA encoding the T<sub>pase</sub> was co-injected with the tester pJKI216 or *shh-GOHS* target and *gfp-fg* donor DNA into 1-cell-stage eggs. Total DNA was extracted from embryos and analyzed. Symbols are as follows: black and white triangles in open boxes represent the left and right ends of IS30; numbered black boxes show exons of *shh* locus (thick line), marker genes are highlighted by open boxes. Abbreviations: Ap<sup>R</sup>, Km<sup>R</sup> and Cm<sup>R</sup> – ampicillin, kanamycin and chloramphenicol resistance genes, respectively; GFP – green fluorescence marker gene; intA – splice acceptor from the first intron of carp  $\beta$ -actin gene; GOHS – hot spot target sequence of IS30; polyA – polyadenylation signal of SV40. The sizes of junction fragments between *shh-GOHS* target and *gfp-fg* donor amplified with primer pairs S3-IRL and cat5'-neo5' in sense orientation and S3-IIdi6 and seq260-neo5' in reverse orientation are indicated with a horizontal bar.

mMessage mMachine kit. Microinjection of 1 or 2 cell stage eggs was carried out as described [14]. T<sub>pase</sub> mRNA was used in 100 ng/ $\mu$ l concentration in combinations with circular *gfp-p* or *gfp-fg* donor and *shh* targets (200 ng/ $\mu$ l each). Zebrafish were kept and embryos generated according to [15]. Injected embryos were grown to 1–3 somites stage and harvested for DNA preparation. 200 embryos were digested in SET buffer overnight at 55°C with 40  $\mu$ g/ml proteinase K and DNA was prepared by phenol extraction [14].

### 3. Results and discussion

#### 3.1. Trans-kingdom activity of IS30 in zebrafish embryos and human cells

First, IS30 T<sub>pase</sub>-mediated excision activity was examined in the vertebrate system, zebrafish embryo. The experimental strategy included the microinjection of a two-component transposition system developed previously for *E. coli* [7] into fish embryos. The T<sub>pase</sub> was expressed from a 'producer plas-

mid' or synthetic mRNA in the presence of the tester plasmid pJKI216 (Fig. 1A). Total DNA extracted from pooled 10 hpf embryos was introduced into *E. coli* and the in trans activity of the T<sub>pase</sub> was monitored by recovery and analysis of the tester plasmid population. Cm<sup>S</sup>Km<sup>R</sup> colonies were determined by replica-plating and plasmid DNA of some colonies were sequenced to assess the correct excision of the Cm<sup>R</sup> gene. Activity of the T<sub>pase</sub> was characterized by measuring the fraction of Cm<sup>S</sup> clones among Km<sup>R</sup> transformants. The Cm<sup>S</sup> frequency was 4–28-fold higher in samples exposed to T<sub>pase</sub> than in controls (tester injected alone), showing that the prokaryotic T<sub>pase</sub> is functional in zebrafish (Table 1). Similar frequencies were obtained using T<sub>pase</sub> producer plasmid or T<sub>pase</sub> mRNA. Joining of the MT or a NLS to IS30 T<sub>pase</sub> did not significantly affect the transposition frequency indicating that the T<sub>pase</sub> can be fused to other peptides without losing its activity.

Table 1  
Transpositional excision mediated by IS30 T<sub>pase</sub> in zebrafish and Hela cells

| Components injected          | Total Km <sup>R</sup> colonies tested | Number of                                |                   |                      | Correct excisions (%) |
|------------------------------|---------------------------------------|--|-------------------|----------------------|-----------------------|
|                              |                                       | Km <sup>R</sup> Cm <sup>S</sup> colonies | Correct excisions | Other recombinations |                       |
| Zebrafish                    |                                       |  |                   |                      |                       |
| tester (pJKI216)             | 1949                                  | 2  | 1 <sup>a</sup>    | 1                    | <b>0.05</b>           |
| tester+Tpase producer        | 1680                                  | 15                                       | 12                | 3                    | <b>0.71</b>           |
| tester+NLS-Tpase producer    | 1718                                  | 14                                       | 10                | 4                    | <b>0.58</b>           |
| tester+NLS-MT-Tpase producer | 516                                   | 4  | 4                 | 0                    | <b>0.78</b>           |
| tester                       | 1848                                  | 1  | 0                 | 1                    | <b>&lt; 0.05</b>      |
| tester+Tpase mRNA            | 907                                   | 3  | 2                 | 1                    | <b>0.22</b>           |
| tester+NLS-Tpase mRNA        | 787                                   | 13                                       | 11                | 2                    | <b>1.40</b>           |
| tester+MT-Tpase mRNA         | 567                                   | 8  | 7                 | 1                    | <b>1.41</b>           |
| tester+NLS-MT-Tpase mRNA     | 282                                   | 3  | 3                 | 0                    | <b>1.06</b>           |
| Hela cells                   |                                       |  |                   |                      |                       |
| tester (pJKI216)             | 3715                                  | 0  | 0                 | 0                    | <b>&lt; 0.026</b>     |
| tester+Tpase producer        | 1722                                  | 2  | 2                 | 0                    | <b>0.12</b>           |

<sup>a</sup>This excision event may represent the background T<sub>pase</sub> activity in *E. coli* [7].

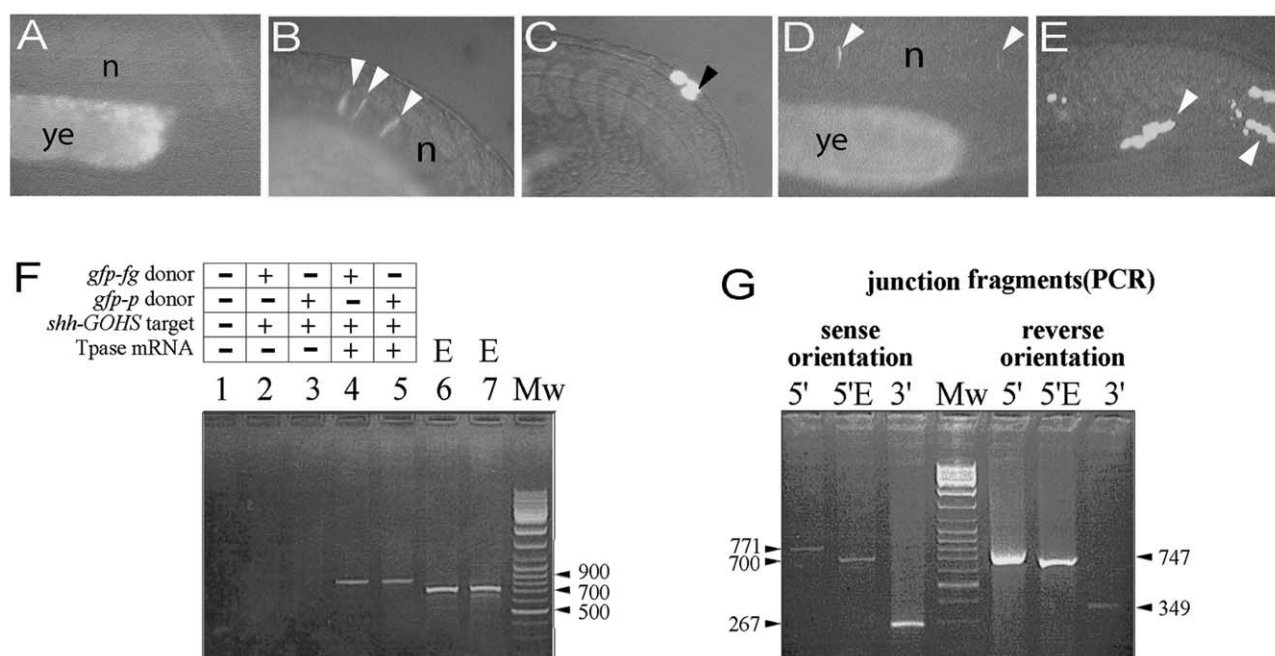


Fig. 2. Insertion of a *gfp* donor into the *shh* target induced by IS30 Tase results in tissue-specific expression of the *gfp* reporter gene in zebrafish embryos. A–E: Detection of GFP expression in one-day-old zebrafish embryos. A: Tail region of a control non-injected embryo. B: Detection of GFP expression in notochord cells of embryos injected with the control *shh-gfp* fusion construct. C: Ectopic GFP activity in cells where *shh* is not expressed (epidermis). D: Embryo injected with *gfp-fg* or *gfp-p* (see Section 2) donors and *shh-GOHS* target along with Tase mRNA expresses GFP in notochord cells suggesting regulation of the *gfp* gene by the *shh* locus. E: Embryo from the same injection showing non-specific activity in epidermal cells. Abbreviations: n – notochord, ye – yolk extension. Embryos are oriented anterior left. Lateral views of the tail region are shown. F: PCR analysis demonstrating the presence of junction fragments between microinjected *gfp-fg* donor and the *shh-GOHS* target plasmid. The experimental setup is presented over the lanes. Lanes 6 and 7 show the verification of PCR products from lanes 4 and 5 by *Eco*RI digestion. G: PCR reaction to detect the junction fragments produced by integration of *gfp-fg* donor into *shh-GOHS* target in sense and reverse orientations. Abbreviations: E – *Eco*RI digested PCR product, Mw – molecular weight marker, 5' and 3' – junction fragments generated at the 5' or the 3' end of inserted donor, respectively.

Capacity of Tase for promoting the excision in human cells was also assessed with the same system. Hela cells were transfected with the tester plasmid alone or along with the Tase producer plasmid. DNA from transfected cells was transformed into *E. coli* and subjected to antibiotic selection and sequencing. Correct excision of the *Cm<sup>R</sup>* gene could not be detected (<0.026%) with the tester plasmid alone, while 0.12% of plasmids were *Cm<sup>S</sup>* when the Tase producer was co-transfected (Table 1). Sequencing the plasmids recovered from the *Cm<sup>S</sup>* colonies confirmed that both recombination events detected were correct transpositional deletions (data not shown). Excisions catalyzed by IS30 Tase require the interaction of two recombinogenic (IS30)<sub>2</sub> structures (Fig. 1A). Conversely, insertion of a donor DNA into a recipient target involves only one (IS30)<sub>2</sub> structure and a target (hot spot) sequence in the recipient DNA (Fig. 1B). To test the ability of Tase to promote insertions in fish embryos, a three-component gene trap system was constructed. The system consists of an (IS30)<sub>2</sub> donor, a 'hot spot' target and the IS30 Tase. The donor DNA contains the *gfp* reporter gene preceded by the splice-acceptor sequence *intA* and the (IS30)<sub>2</sub> recombination site. The promoterless *gfp* gene is expected to be inactive unless integration of this construct occurs into an intronic sequence of a transcribed gene. Donor DNA was there-circularized *gfp-fg* fragment containing each functional part for gene trapping (Fig. 1B, see Section 2) but lacking the plasmid backbone. In the target plasmid *shh-GOHS*, the artificial IS30 hot spot sequence GOHS generated according to the consensus of *E. coli* genomic IS30 integration sites [11]

was inserted into the first intron of the *sonic hedgehog* (*shh*) gene of zebrafish. This construct contained also the transcriptional regulatory elements driving expression of *shh* in the notochord and the ventral neural tube of zebrafish embryos [16]. When correctly oriented, the integration of *gfp-fg* donor into GOHS will result in *gfp* expression in the midline of neurula and organogenesis stage embryos.

To assess whether a sense *shh-gfp* fusion can express *gfp* faithfully, a fusion plasmid generated by transposition in *E. coli* was injected into embryos and *gfp* expression was analyzed at 24 hpf. Mosaic green fluorescent protein (GFP) activity restricted to several cells per embryo due to uneven segregation of the foreign DNA [17]. Expression of *gfp* in notochord cells was observed in eight out of 30 embryos (26 notochord cells out of 65 *gfp*-positive cells), indicating that *gfp* was expressed under the control of regulatory elements of the *shh* gene (Fig. 2B). Non-specific expression was also detected in ectodermal and other cell types, which is explained by cryptic non-specific regulatory elements in the carp  $\beta$ -actin intronic sequences adjacent to the splice-acceptor site (Fig. 2C).

To detect integration events in zebrafish, the *gfp-fg* donor and the *shh-GOHS* target DNA were injected into embryos with or without synthetic mRNA encoding the Tase. Notochord cells expressing *gfp* were detected in four out of 130 embryos at 24 hpf (Fig. 2D), while none of the embryos (0/195 embryos) showed notochord-specific *gfp* expression when Tase mRNA was omitted (Fig. 2E) or when the *gfp-fg* was injected alone (data not shown). This suggests that the inser-



tion into *shh-GOHS* target was induced by the T<sub>pase</sub>. To verify this conclusion, total DNA was prepared from microinjected embryos pooled together and the junction fragments of *shh-gfp* fusion were PCR amplified (Fig. 1B). Correct integration of the *gfp* cassette into the *shh-GOHS* target was obtained only when T<sub>pase</sub> mRNA was coinjected (Fig. 2F). The junction fragments were confirmed by restriction analysis and sequencing (Fig. 2G). As expected, insertions occurred in both orientations. Frequency of integration was estimated by transformation of DNA isolated from microinjected embryos into *E. coli*. The frequency of *shh-gfp* fusion plasmids was determined as a ratio of  $\text{Cm}^R\text{Ap}^R/\text{Cm}^S\text{Ap}^R$  transformants, which was  $5.1 \times 10^{-5}$  and  $5.8 \times 10^{-5}$  in two independent experiments, respectively. The correct structure of *shh-gfp* fusions was verified by sequencing. This is the first demonstration that a prokaryotic IS element is functional in a vertebrate system. The activity of IS30 T<sub>pase</sub> in fish cells may indicate a theoretical possibility for horizontal gene transfer via transposon-mediated recombination between prokaryotic and eukaryotic genomes.

### 3.2. Modification of targeting specificity of the IS30 T<sub>pase</sub>

We tested in *E. coli* whether the target specificity of IS30 T<sub>pase</sub> can be altered by fusing it to another DNA binding protein. The T<sub>pase</sub> was attached to the cI repressor of bacteriophage  $\lambda$ . The activity of the chimeric protein was examined in a two-component transposition system consisting of an (IS30)<sub>2</sub> donor plasmid that also expressed the T<sub>pase</sub>-cI fusion protein and a target plasmid carrying the  $\lambda$  operator  $O_R$  (Fig. 3A). The chimeric T<sub>pase</sub> retained the activities of both domains as suggested by the immunity of *E. coli* carrying the fusion protein against  $\lambda$  infection, and by the similar activity as compared to wt T<sub>pase</sub> in promoting insertions into the GOHS hot spot ( $2.5 \times 10^{-2}$  and  $2.0 \times 10^{-2}$  for wt and chimera T<sub>pase</sub>, respectively). Next, the capability of the chimeric T<sub>pase</sub> to direct the insertions into the  $\lambda$  operator  $O_R$  was investigated. The overall frequency of transposition was  $1.5 \times 10^{-4}$ . Among 53 transposition products, 30 carried the inserted donor plasmid in close proximity (within 400 bp) of  $O_R$  and one insertion occurred directly in the  $O_R$  (Fig. 3B). Additional 22 insertions were farther from the  $O_R$ . These insertions depended also on the cI repressor domain in the chimera as no integration into the  $O_R$  target plasmid was scored when the wt T<sub>pase</sub> was used ( $< 2.5 \times 10^{-5}$ ). Moreover, deletion of  $O_R$  from the target plasmid abolished its targeting by the T<sub>pase</sub>-cI chimera ( $< 1.5 \times 10^{-5}$ ). Seventeen sequenced insertion sites showed no or limited homology to either the  $\lambda$  operator or to the GOHS sequence (Fig. 3C) suggesting that the chimera has acquired a novel target specificity. Taken together, these results demonstrate that fusion of the cI repressor changes the target selection of IS30 T<sub>pase</sub> and increases the efficiency of integration into target sites not preferred by the wt T<sub>pase</sub> by more than 10-fold in *E. coli*.

We finally assessed whether the target specificity of T<sub>pase</sub> can also be directed to new sites in zebrafish embryos by fusion with a vertebrate DBD. The three-component strategy described in Fig. 1B was applied with modifications. The DBD of the transcription factor Gli1 [12] was fused to the C-terminus of IS30 T<sub>pase</sub> (T<sub>pase</sub>-GliDBD). We employed *shh-gli* plasmid as target, which carries a Gli1 binding site (GACCACCCA) inserted into the first intron of the *shh* locus. Synthetic mRNA of T<sub>pase</sub>-GliDBD, *gfp-fg* donor and the

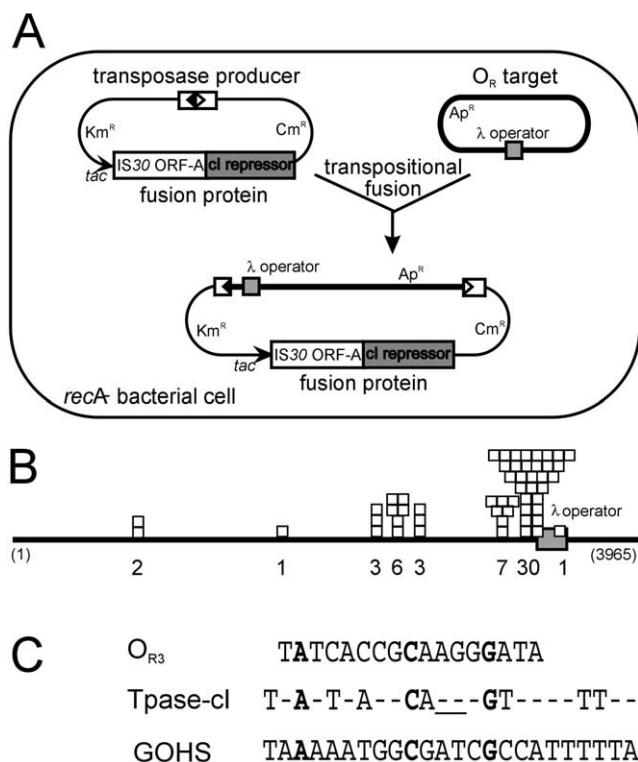


Fig. 3. A: The transposition system based on IS30 T<sub>pase</sub>-cI repressor fusion protein. RecA<sup>-</sup> *E. coli* cells were co-transformed with  $\text{Cm}^R\text{Km}^R$  plasmids coding for T<sub>pase</sub>-cI fusion or wt T<sub>pase</sub> in combination with one of the  $\text{Ap}^R$  target plasmids lacking target site or carrying  $\lambda$  operator  $O_R$  or GOHS hot spot. The transposition assay was carried out as described in [2]. B: Distribution of insertions in  $O_R$  target plasmid. Open boxes represent the sites and the number of integration events detected in the vicinity of the  $\lambda$  operator. C: Comparison of the consensus sequence (more than 50% identity in 17 sequenced integration sites) of the sites targeted by T<sub>pase</sub>-cI fusion to the GOHS hot spot of IS30 and the  $\lambda$   $O_{R3}$  operator. Identical bases in all three sequences are in bold and the integration site in T<sub>pase</sub>-cI consensus target and GOHS is underlined. Symbols are as in Fig. 1; *tac*: *tac* promoter.

*shh-gli* target DNA were coinjected into 1-cell-stage zebrafish embryos. In controls, the injections were performed without T<sub>pase</sub> mRNA or with the *shh* target lacking the Gli1 binding site. Transposition by T<sub>pase</sub>-GliDBD was first analyzed by monitoring *gfp* expression in injected embryos. *Gfp* expression was detected only when all three components of the transposition system were coinjected (10 notochord cells expressed in three out of 104 embryos). No expression in the notochord was seen when *IS30-GliDBD* mRNA and *gfp-fg* donor were coinjected with a *shh* target without the Gli1 binding site (0/102 embryos) or when the *gfp-fg* donor was coinjected with *shh-gli* target in the absence of *IS30-GliDBD* (0/178 embryos). These results indicate that the T<sub>pase</sub>-GliDBD fusion protein is able to direct insertions in the sense orientation into the *shh* target in a Gli1 binding site-dependent manner. To analyze the transposition events further, total DNA was extracted from gastrula-stage embryos and integration of the *gfp* cassette was monitored by nested PCR using primers that detect insertion of *gfp* in the sense orientation within 675 bp upstream and up to 2.5 kb downstream of the Gli1 binding site (Fig. 4). PCR-amplified putative junction fragments were subcloned and sequenced. Fourteen fragments representing 12 different integration events were identified out of which eight

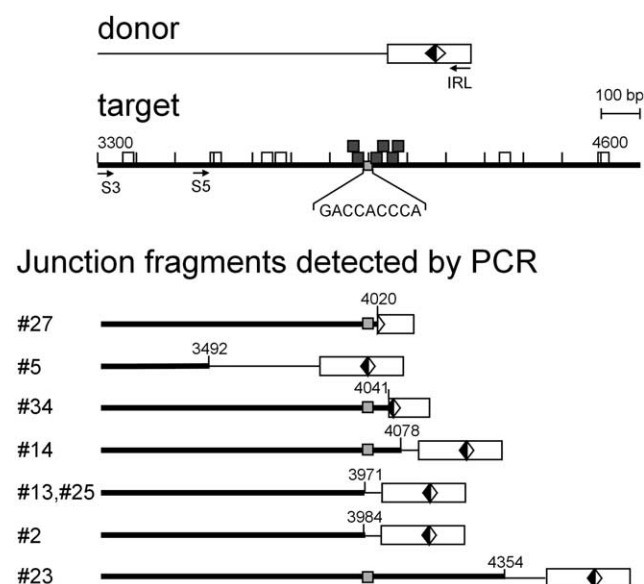


Fig. 4. Directed integration in zebrafish embryos. Analysis of the junction fragments of the *gfp-fg* donor and the *shh-gli* target plasmid detected by PCR. Thin and thick lines represent donor and target sequences, respectively. The light gray box indicates the Gli1 binding site. Squares indicate integration sites detected by PCR. Integrations within 100 bases nearby the Gli1 binding site are highlighted in dark gray. Numbers above the junction fragments show the position of integration of IS-end-containing donor sequence in the *shh-gli* target DNA. Black arrows indicate the annealing positions of S3, S5 and IRL oligonucleotides used in PCR reactions. Note that this strategy detects insertion in the sense orientation only. Other symbols are the same as in Fig. 1.

are shown in Fig. 4. In two cases the same fragment was identified twice (e.g. clone #13 and #25). The resolution of (IS30)<sub>2</sub> indicative of a legitimate transposition event was detected only in one case (clone #27). This integration occurred 36 bp away from the 3' end of the Gli1 binding site. In total, six integrations were located within 100 bases adjacent to the Gli1 binding site. In most integration, however, the (IS30)<sub>2</sub> site was intact, suggesting that the insertions took place by illegitimate recombination. No junction fragments could be amplified when the Gli1 binding site was missing in the target plasmid. These results suggest that the level of both transpositional as well as illegitimate recombination in the vicinity of the Gli1 binding site is dependent on the chimeric Tase. Given the large variety of available DNA binding specificities, this feature of IS30 Tase to be modulated by a linked DBD can overcome the known bias in target site selection of the wt protein and allows for integration in a vast variety of different sites. The efficiency of both excision and integration was relatively low, however, similar frequencies were observed in *E. coli* [1,2].

It was shown previously, that the integrase of HIV-1 can be tethered to predetermined location by the fusion to the DBD of the  $\lambda$  cI repressor in an in vitro system [18,19]. Our results with the IS30 Tase in *E. coli* and zebrafish embryos indicates that targeting of transposition may be generally achieved in living cells. An integration system that can be targeted to predetermined positions on DNA may have numerous applications. Such a system could be utilized, for example, to search for and to tag the targets of DNA binding proteins in vivo. Moreover, the integration of donor sequences targeted by a DBD of a transcription factor may cause insertion

mutation by disrupting the regulatory elements of a target gene of the transcription factor. In addition to the promising features demonstrated here, IS30 Tase has a number of characteristics that may render the Tase as tool for functional genomics worthwhile. These include the lack of size limitations, the high activity of the (IS30)<sub>2</sub> structure that is resolved into a much less active integration product and the absence of homologues of IS30 Tase in eukaryotic cells. Especially the latter features will guarantee that the transposition products are stable once integrated. A central issue of any further development is to improve the frequency of recombination. Fusion to a heterologous DBD did not only alter the range of target sequences but increased also the frequency of integration in both *E. coli* and zebrafish. This suggests that the frequency may simply be increased by more efficient tethering the transposase to DNA. Fusion of transposases with general transcription factors such as the TATA box binding factors or with DNA methylases may also improve the efficiency of integration into a vast array of potential integration sites.

**Acknowledgements:** We are grateful to Á. Bakos, I. Keresztúri, E. Keresztúri, Zs. Lengyel, M. Turai, N. Fischer for skillful technical assistance. We thank L. Tora for useful suggestions. This study was supported by grants F6/99 of Tét Foundation, NKFP4/006 of Ministry of Education (to F.O.) and of Hungarian National Scientific Research Fund (OTKA) T019365, T025729 (to F.O.) and F038204 (to J.K.). U.S. was supported by INSERM, CNRS, HUS, ACI, VW-Stiftung, AFM, ARC and AICR. F.M. was supported by AFM and EMBO.

## References

- [1] Kiss, J. and Olasz, F. (1999) Mol. Microbiol. 34, 37–52.
- [2] Olasz, F., Stalder, R. and Arber, W. (1993) Mol. Gen. Genet. 239, 177–187.
- [3] Ivics, Z., Hackett, P.B., Plasterk, R.H. and Izsvak, Z. (1997) Cell 91, 501–510.
- [4] Koga, A., Suzuki, M., Inagaki, H., Bessho, Y. and Hori, H. (1996) Nature 383, 30.
- [5] Gueiros-Filho, F.J. and Beverley, S.M. (1997) Science 276, 1716–1719.
- [6] Olasz, F., Kiss, J., König, P., Buzas, Z., Stalder, R. and Arber, W. (1998) Mol. Microbiol. 28, 691–704.
- [7] Farkas, T., Kiss, J. and Olasz, F. (1996) FEBS Lett. 390, 53–58.
- [8] Turner, D.L. and Weintraub, H. (1994) Genes Dev. 8, 1434–1447.
- [9] Alam, M.S., Lavender, F.L., Iyengar, A., Rahman, M.A., Ayad, H.H., Lathe, R., Morley, S.D. and Maclean, N. (1996) Mol. Reprod. Dev. 45, 117–122.
- [10] Chang, B.E., Blader, P., Fischer, N., Ingham, P.W. and Strahle, U. (1997) EMBO J. 16, 3955–3964.
- [11] Olasz, F., Kiss, J., König, P., Buzas, Z., Stalder, R. and Arber, W. (1998) Mol. Microbiol. 28, 691–704.
- [12] Kinzler, K.W. and Vogelstein, B. (1990) Mol. Cell. Biol. 10, 634–642.
- [13] Sambrook, J. and Russell, D.W. (2001) Molecular cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [14] Muller, F., Williams, D.W., Kobolak, J., Gauvry, L., Goldspink, G., Orban, L. and Maclean, N. (1997) Mol. Reprod. Dev. 47, 404–412.
- [15] Westerfield, M. (1995) The Zebrafish Book, University of Oregon Press, Eugene, OR.
- [16] Muller, F., Chang, B., Albert, S., Fischer, N., Tora, L. and Strahle, U. (1999) Development 126, 2103–2116.
- [17] Iyengar, A., Muller, F. and Maclean, N. (1996) Transgenic Res. 5, 147–166.
- [18] Bushman, F.D. (1994) Proc. Natl. Acad. Sci. USA 91, 9233–9237.
- [19] Bushman, F.D. and Miller, M.D. (1997) J. Virol. 71, 458–464.